

Artificial Intelligence Augmentation of Radiologist Performance in Distinguishing COVID-19 from Pneumonia of Other Origin at Chest CT

Year: 2021 | Citations: 301 | Authors: H. Bai, Robin Wang, Z. Xiong, B. Hsieh, Ken Chang

Abstract

Background Coronavirus disease 2019 (COVID-19) and pneumonia of other diseases share similar CT characteristics, which contributes to the challenges in differentiating them with high accuracy. **Purpose** To establish and evaluate an artificial intelligence (AI) system for differentiating COVID-19 and other pneumonia at chest CT and assessing radiologist performance without and with AI assistance. **Materials and Methods** A total of 521 patients with positive reverse transcription polymerase chain reaction results for COVID-19 and abnormal chest CT findings were retrospectively identified from 10 hospitals from January 2020 to April 2020. A total of 665 patients with non-COVID-19 pneumonia and definite evidence of pneumonia at chest CT were retrospectively selected from three hospitals between 2017 and 2019. To classify COVID-19 versus other pneumonia for each patient, abnormal CT slices were input into the EfficientNet B4 deep neural network architecture after lung segmentation, followed by a two-layer fully connected neural network to pool slices together. The final cohort of 1186 patients (132 583 CT slices) was divided into training, validation, and test sets in a 7:2:1 and equal ratio. Independent testing was performed by evaluating model performance in separate hospitals. Studies were blindly reviewed by six radiologists without and then with AI assistance. **Results** The final model achieved a test accuracy of 96% (95% confidence interval [CI]: 90%, 98%), a sensitivity of 95% (95% CI: 83%, 100%), and a specificity of 96% (95% CI: 88%, 99%) with area under the receiver operating characteristic curve of 0.95 and area under the precision-recall curve of 0.90. On independent testing, this model achieved an accuracy of 87% (95% CI: 82%, 90%), a sensitivity of 89% (95% CI: 81%, 94%), and a specificity of 86% (95% CI: 80%, 90%) with area under the receiver operating characteristic curve of 0.90 and area under the precision-recall curve of 0.87. Assisted by the probabilities of the model, the radiologists achieved a higher average test accuracy (90% vs 85%, $\Delta = 5$, $P < .001$), sensitivity (88% vs 79%, $\Delta = 9$, $P < .001$), and specificity (91% vs 88%, $\Delta = 3$, $P = .001$). **Conclusion** Artificial intelligence assistance improved radiologists' performance in distinguishing coronavirus disease 2019 pneumonia from non-coronavirus disease 2019 pneumonia at chest CT. © RSNA, 2020 Online supplemental material is available for this article.